IHS National Data Warehouse

Tech Fair Presentations

August 2004
Office of Information Technology
Albuquerque, NM

Project Manager:	Stanley Griffith
Deputy Project Managers:	Joseph Herrera
	Michael Gomez
Contractor Leads:	Paul Golis
	Lisa Petrakos

1

National Data Warehouse

An Overview for Clinicians and Managers

- Why Do We Need National-level Data?
- Why a DW?
- DW Design
- User Benefits
- Timeline
- The Future

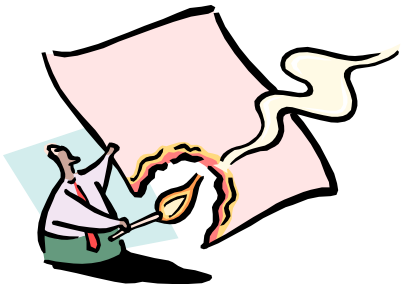
6

Why Do We Need National Data?

- Requests for information are ever increasing
- These ever increasing requests will likely continue
- But our resources will not increase commensurate to these increased demands

7

Why Do We Need National Data?



- By deriving information from existing data, data that are already being collected for other purposes,

  we can reduce local data collection burdens and bureaucratic overhead at all levels

8

Why Do We Need National Data?

Information means money

- GPRA is tied to budget requests
- Funding for special needs
  - Racial disparities
  - Diabetes
- ORYX and JCAHO accreditation

9

Why Do We Need National Data?

National level data can improve clinical care

- Patient safety

10

Why Do We Need National Data?

Information provides program direction

- Needs assessment
- Resource allocation
- Outcome performance
- Facility planning

11

National Data Warehouse
An Overview for Clinicians and Managers

- Why Do We Need National-level Data?
- Why a DW?
- DW Design
- User Benefits
- Timeline
- The Future

12

What kind of a repository do we need?

- We have conflicting needs.

13

We need more data…

The breadth of information needs is increasing

- Statistical
- ORYX
- Epidemiology
- Diabetes
- Pharmacy
- Cost management
- Bioterrorism

etc.

14

… in one, all-inclusive repository…

One central collection point for all data

- Verify integrity of data
- Analyze and provide feedback on
  - Timeliness of data
  - Data quality
  - Unexpected deviations from historical norms
- Maintain a "single version of historical truth"
- Maintain all the information content of the data

15

… but …

- One large information repository becomes just too unwieldy for reporting

16

… we need focused, efficient DBs…

We need efficient, user friendly access to data

- Ease of user access
- Search-efficient DB structures
- Subsets of data – just what we need for specific uses
- Transformed data

17

… still, on the other hand…

But data marts alone just cannot maintain the

- Flexibility
- Granularity
- Scalability
- Power

that will be required to meet all needs now or in the future

18

… because we have to accommodate the future.

The architecture has to be able to accommodate change, inevitable change.

There will be future needs for information that even the most astute among us cannot now anticipate.

19

Why Do We Need National Data?

- We need one, large, replete, and powerful *data warehouse database*,
- that provides information to more focused, user-friendly, search-efficient *data marts*.
- We need a ***data warehouse environment***.

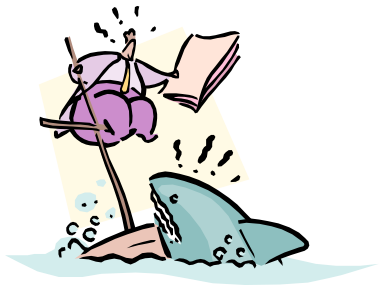
20

National Data Warehouse
An Overview for Clinicians and Managers

- Why Do We Need National-level Data?
- Why a DW?
- DW Design
- User Benefits
- Timeline
- The Future

21

The Dilemma…



- We need a complete, all-inclusive solution and we needed it yesterday
- But to plan, build, and implement a complete solution would take far, far too long and would not deliver incremental value to our users along the way

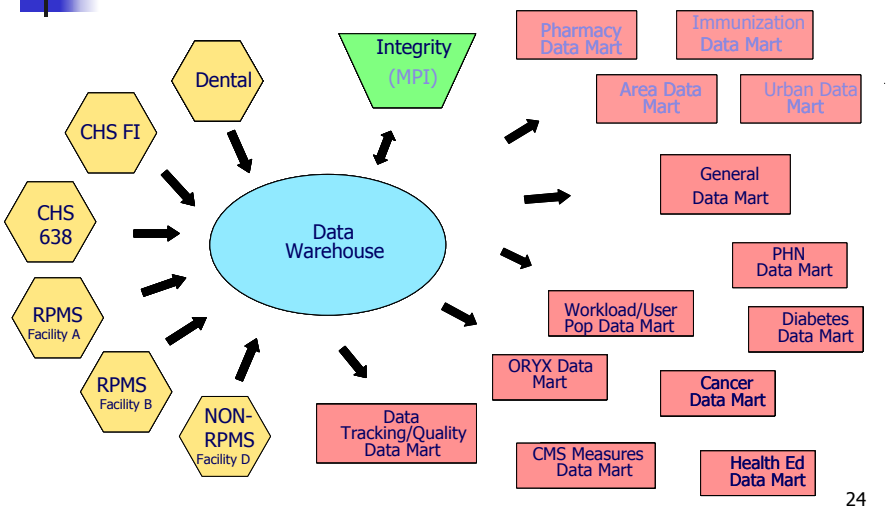
22

The Solution

- Build the data warehouse environment in increments that deliver value to users at each step
- The most critical user needs will be addressed first

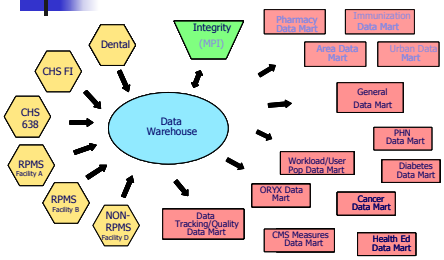
23

High-Level NDW Architecture



24

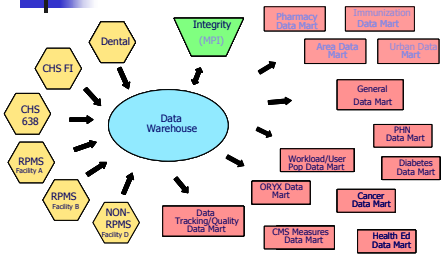
The DW1 Environment Will…



- Accept registration and encounter-based patient data…
- …from sites using RPMS or other IT systems…
- …and CHS data from the Fiscal Intermediary and Area databases.

25

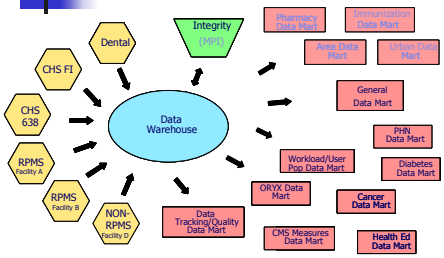
The DW1 Environment Will…



- Employ healthcare industry standards (HL7) for data transport and messaging …
- …employing an Integration engine to route those data from sites to the proper destinations (e.g., the DW, an Area database).

26

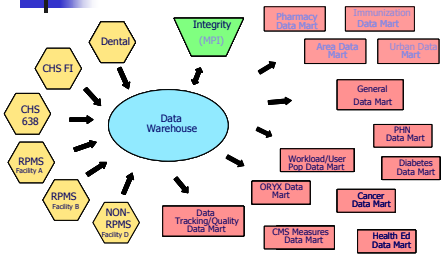
The DW1 Environment Will…



- Collect and store information almost exactly as it is received from the field …
- …maintaining historical snapshots of those data…
- …maintaining as much of its informational content, as possible.

27

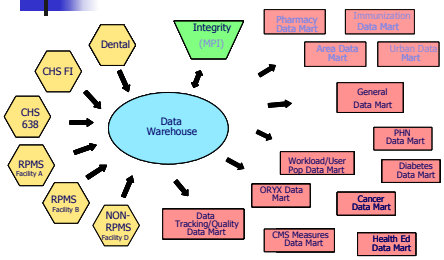
The DW1 Environment Will…



- Provide accurate and helpful information about the data it receives to our field colleagues:
  - Record counts
  - Timeliness of data
  - Less than expected counts based on historical benchmarks
  - Missing data in fields, erroneous codes

28

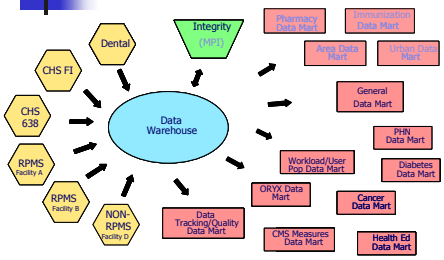
The DW1 Environment Will…



- Supply data to various data marts from which more targeted information can be gathered
  - Workload/User Pop
  - Data Tracking/Quality
  - Outcome measurement (ORYX, GPRA)
  - Organizational units (Area, Urban)
  - Programs (diabetes, cancer, health education, public health nursing, epidemiology, dental, pharmacy)

29

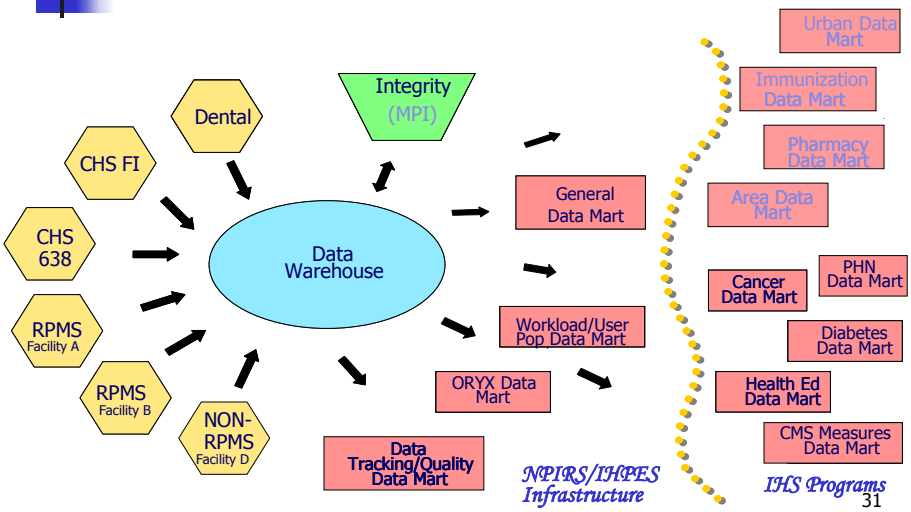
The DW1 Environment Will…



- Use a "probabilistic matching application to resolve by individual across facilities…
- …and use a more complete "Master Person Index" application, when available.

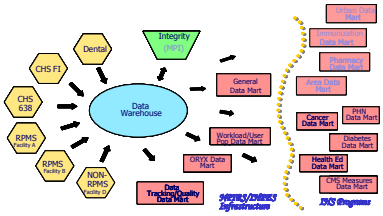
30

NDW Organizational Architecture



31

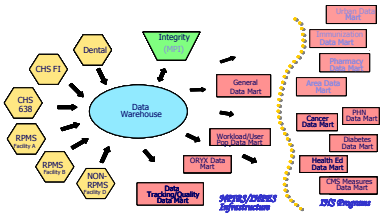
NDW Organizational Architecture



- Underlying DW database, technical infrastructure, general user data marts
  - Primary responsibility – ITSC/IHPES
  - ETL, staging and target tables, administrative and lookup tables
  - "Super users" have access to general data mart that contains a wide spectrum of current data
  - ORYX, CMS Measures, and User Pop/WL data marts
  - Provides data to populate the the more targeted data marts
  - Users of this component are the "super-users" of the general data marts and the more targeted data marts

32

NDW Organizational Architecture



- "Component-ized" targeted data marts
  - Primary responsibility – programs, disciplines, organizational units
  - ITSC "hosts" the infrastructure, provides the security protocols, etc.
  - Very specific, limited subsets of data
  - Emphasis on search efficiency
  - May pull data from general data marts rather than DW database itself
  - Main focus are pre-structured reports
  - Users of this component are the programs, disciplines, organizational units

33

National Data Warehouse
An Overview for Clinicians and Managers

- Why Do We Need National-level Data?
- Why a DW?
- DW Design
- User Benefits
- Timeline
- The Future

34

Benefits for our Users

- Improved report accuracy because of improved linking between registration and encounter records
- More confidence in data - better able to verify reports with local data
- Improved data tracking and data quality assessment tools
- Better able to report patient "outcomes"
- Better support for Areas' performance measurements, health status assessments, and surveillance

35

Benefits for our Users (cont.)

- Better consistency among reports - a rigorously gathered, verified, and maintained "single version of truth"
- Lessened local burdens in producing GPRA measures
- More choices for ORYX users
- More information available for Tribal EPI Centers
- Better information to support Diabetes monies

36

National Data Warehouse
An Overview for Clinicians and Managers

- Why Do We Need National-level Data?
- Why a DW?
- DW Design
- User Benefits
- Timeline
- The Future

37

DW1 Timeline

- Began final initial data load on April 2004
- Integrated step-wise loading and testing over 9-10 months
- Include historical load of data (encounters dated October 1, 2000 to present)
- Run in parallel with current production systems (NPIRS & ORYX) until testing complete
- Legacy systems will be used for
  - FY 2004 WL/UP reports
  - ORYX reports through the April 30, 2005 report
  - All other production reports through Spring 2005
- NDW1 will be used for
  - FY 2005 WL/UP reports
  - July 31, 2005 ORYX reports
  - Other production reports starting mid summer 2005

38

National Data Warehouse
An Overview for Clinicians and Managers

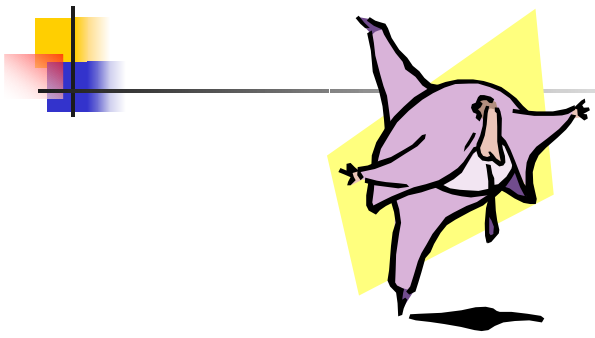
- Why Do We Need National-level Data?
- Why a DW?
- DW Design
- User Benefits
- Timeline
- The Future

39

Future Directions

- Many, many *more data marts*
- *More detailed encounter data* (e.g., labs, meds, measurements, patient education)
- Data from *additional sources* (e.g., Vital Statistics, CMS, Renal Network)
- *New types of patient-related clinical data* (problem list, surgical history, allergies)
- *Other types of data* (claims, purchase orders, community, geographic locations, employee)
- *Better unduplication* of related data from more than one source
- *Enhanced data quality, error, and tracking reports*
- *Archiving* system

40



Questions?

Please visit our web site at:
www.ndw.ihs.gov

41